StackOverflow Developer Survey IBM Data Analytics Capstone

Nic Whitehead

July 2022

IBM Developer

SKILLS NETWORK

OUTLINE

- Executive Summary
- Introduction
- Methodology
- Results
  - Visualization – Charts
  - Dashboard
- Discussion
  - Findings & Implications
- Conclusion
- Appendix

EXECUTIVE SUMMARY

- The following slides summarize key findings from an analysis of data collected as part of the 2019 Stack Overflow Developer Survey
- The analysis yielded insights regarding the following:
  - Most popular languages, databases, and other technologies (at the time of data collection)
  - Insight into technologies which are likely to become popular in the future
  - Respondent demographics (e.g. age, gender and location variances among respondents)
- These findings are relevant particularly to current and aspiring developers, recruiters, educators, policy makers and technology leaders within Data-Driven organisations
- It is important to identify future skill requirements and trends to keep pace with changing technologies and remain competitive.

INTRODUCTION

- This presentation has been created for stakeholders and business decision makers within the global IT and business consulting services
- As well as developers, HR professionals, educators and policy makers

- Since 2011, the online programming knowledge sharing platform, Stack Overflow, has been conducting the Stack Overflow Annual Developer Survey
- The objective of the annual survey is to gather data regarding technology usage and trends among developers.
- In this analysis, a subset of the 2019 dataset was examined (present dataset: N = 11,398; original dataset N ≈ 90,000)
- The presentation will help identify future skill requirements in the global IT sector necessary to keep pace with changing technologies and remain competitive

METHODOLOGY

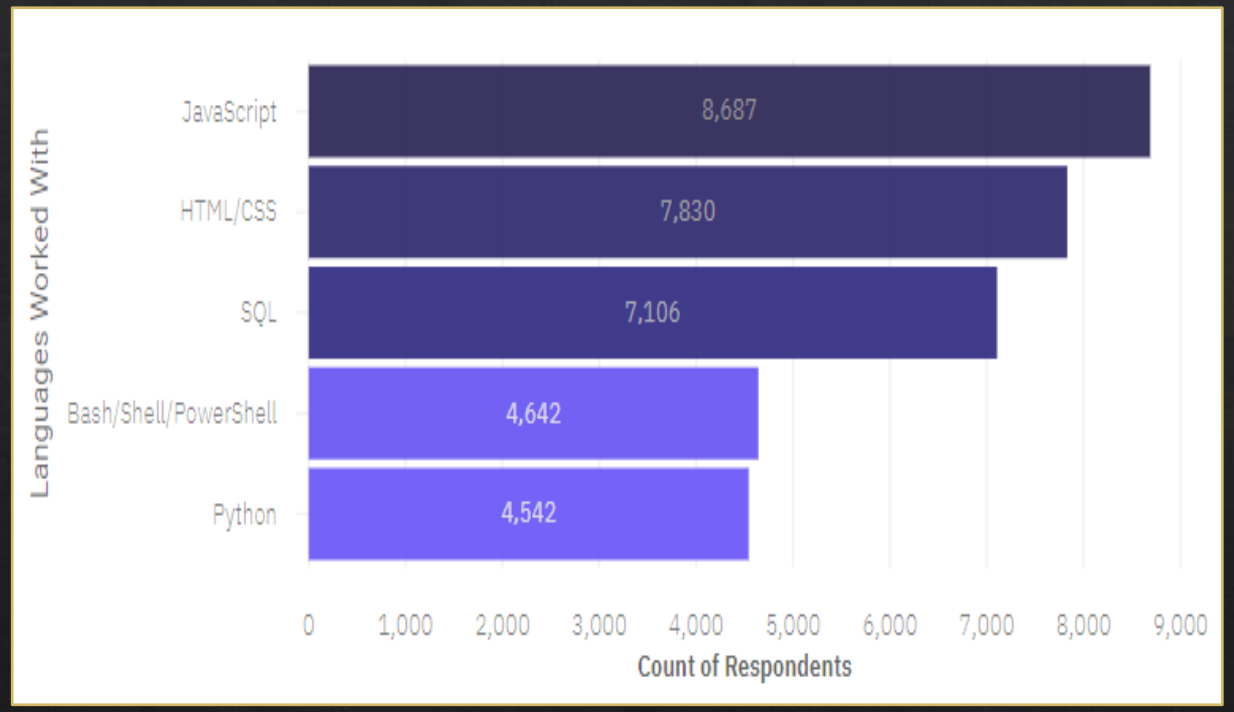
- 2019 Stack Overflow Developer Survey was the raw data source. Link to Stack Overflow annual survey data and results (here)
- A portion of the dataset (provided by IBM) was loaded and cleaned using SQL and Python pandas library
- Duplicates removed, data imputation carried out and data normalization occurred
- Exploratory data analysis and data visualization were conducted using Python libraries and COGNOS
- The following measures were examined:
  - Technologies used in 2019 (languages, databases, platforms, web frames)
  - Technologies most desired for next year
  - Demographics (gender, country, age and education)

RESULTS

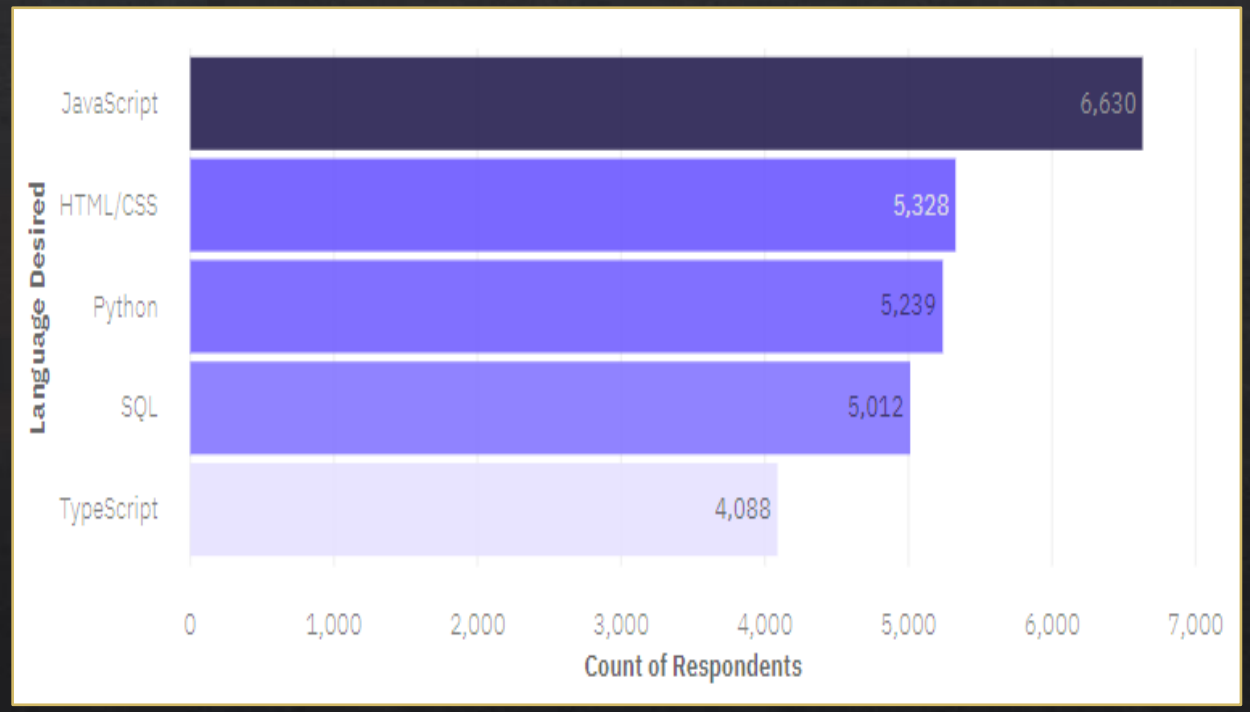
- Descriptive statistics regarding demographics and total annual compensation were computed
- Sample size after compensation outlier removal: N = 10,519 (vs. before removal: N = 11,398)
- The respondents had a median age of 29, and were predominantly male (i.e., 93.5% male vs. 6.5% female)
- Median compensation: $52,704 USD per year
- Median compensation was higher for women: $54,956 than men: $52,339

PROGRAMMING LANGUAGE TRENDS

Current Year



Next Year



IBM Developer

SKILLS NETWORK

PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

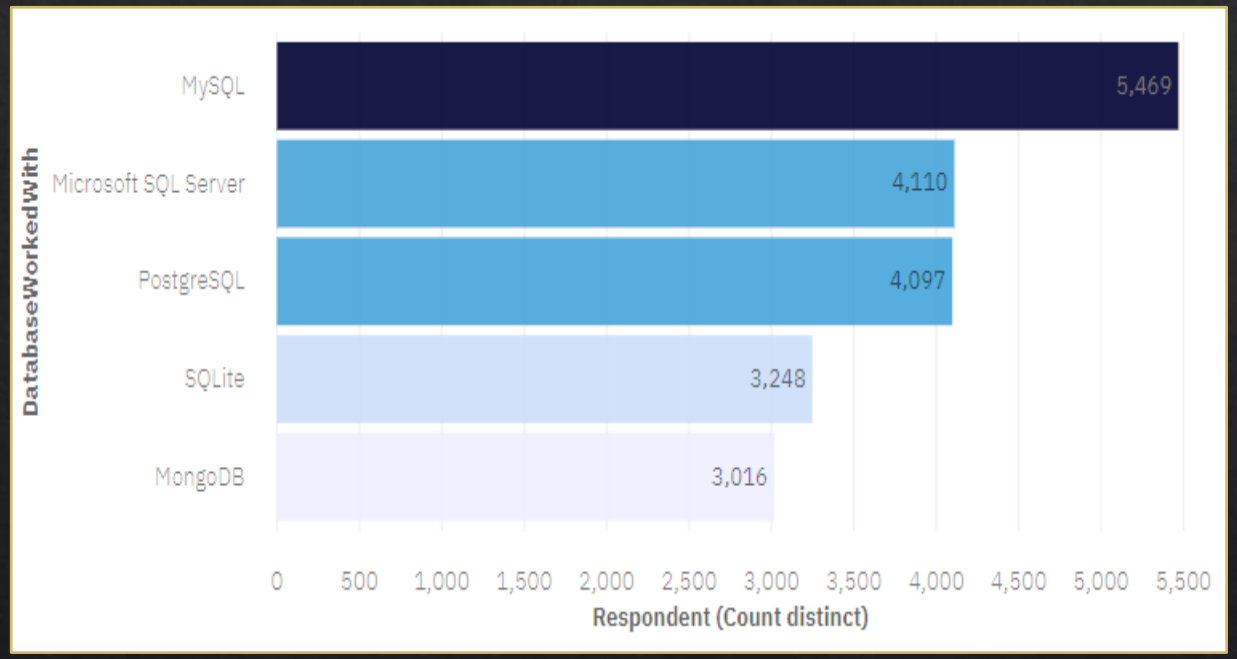
- Javascript and HTML/CSS are currently the two most popular languages in the current year and for next year
- Python and Typescript have seen an increase in popularity from current year to next
- While SQL and Bash/Shell/Powershell have diminished in popularity among respondents

Implications

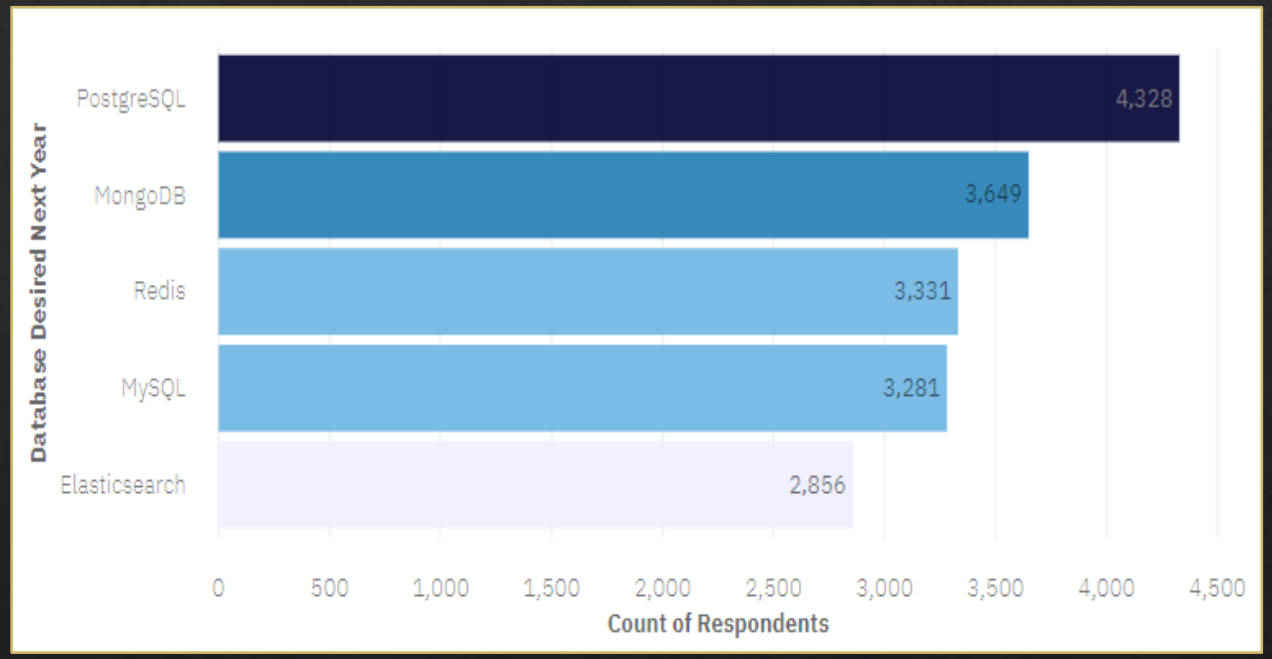
- Businesses should continue to support and employ staff with Javascript and HTML/CSS skills
- Businesses will see an increase in projects and actions performed using Python and Typescript.
- It would be useful to gradually reduce business dependency on applications or models that require SQL or Bash/Shell/Powershell, as resource to maintain these will decrease over time

DATABASE TRENDS

Current Year



Next Year



IBM Developer

SKILLS NETWORK

DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

MongoDB and PostgreSQL have seen a substantial increase in demand for next year

There is also in increase in the demand for Elastisearch and Redis

MySQL, Microsoft SQL Server and SQLite are decreasing in popularity in the next year

Implications

Business should look to provide support and invest in projects that utilise MongoDB and PostgreSQL. There will be increased resource available for these products

Further investigation is suggested in applying Elastisearch and Redis databases. These are seeing an increased level of interest next year.

It would be useful to gradually reduce business dependency on MySQL, MS SQL Server and SQLite databases as resource to maintain these will decrease over time

DASHBOARD

The following link contains the full COGNOS dashboard summarizing current technology use, future technology trend, and demographics of the survey respondents

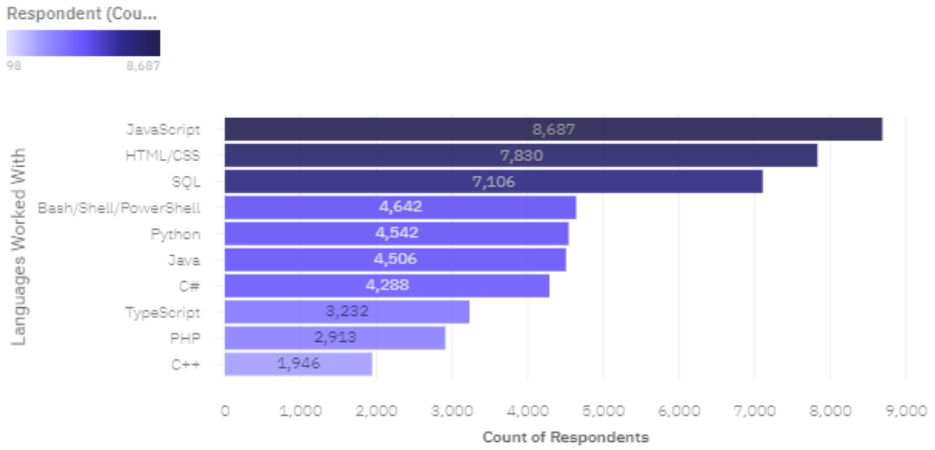
Click here to view the COGNOS dashboard

Screenshots of the dashboard are shown in the next slides

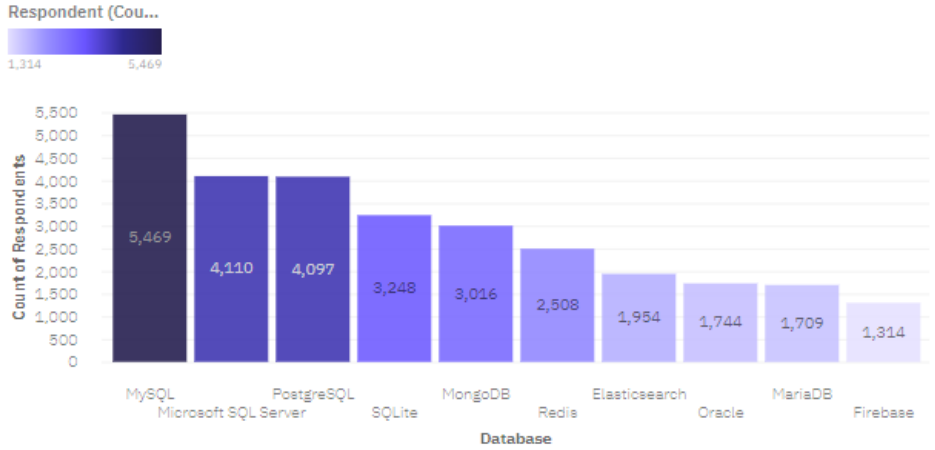
DASHBOARD TAB 1

Current Technology Usage | Future Technology Trend | Demographics

Top 10 Languages Used by Respondents

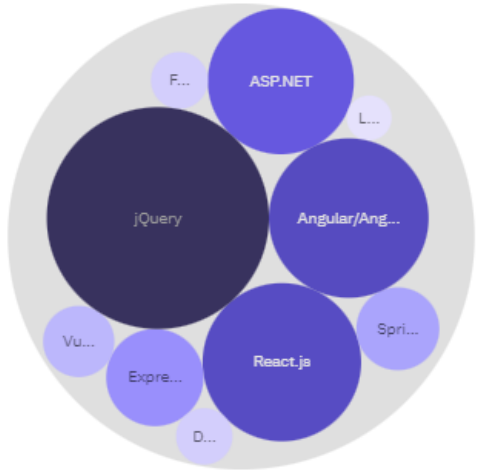


Top 10 Databases Worked with by Respondents



Platforms Worked With

Webframe Worked With

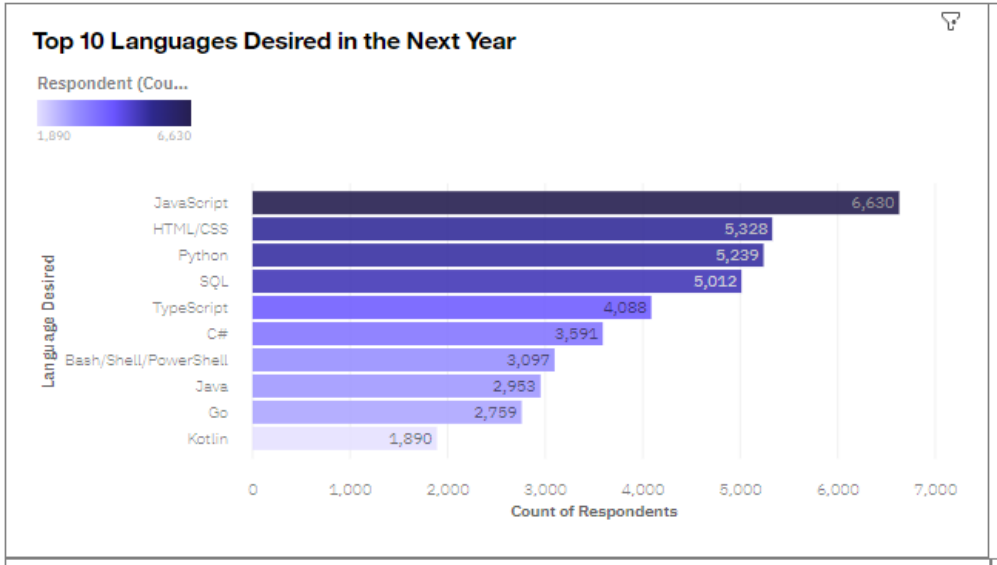


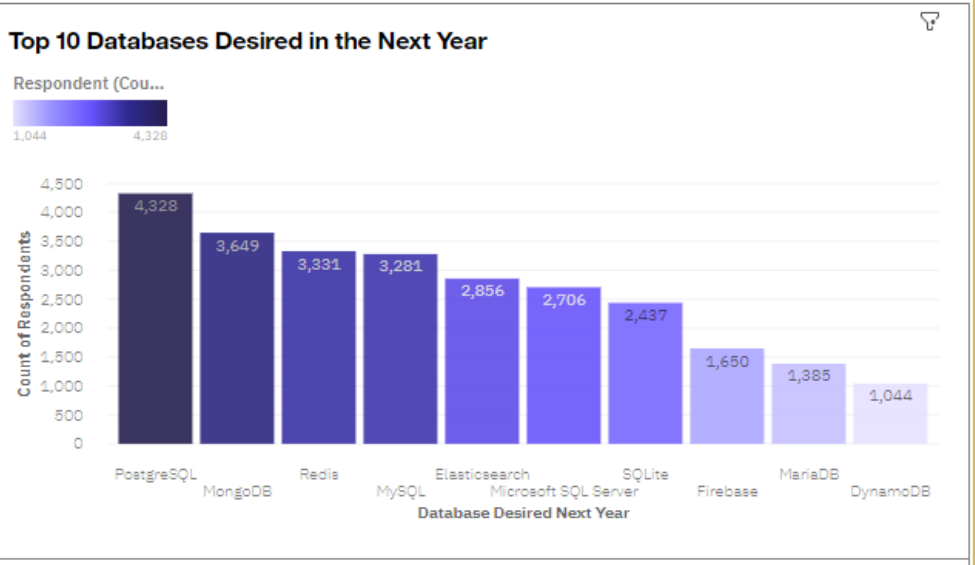
IBM Developer

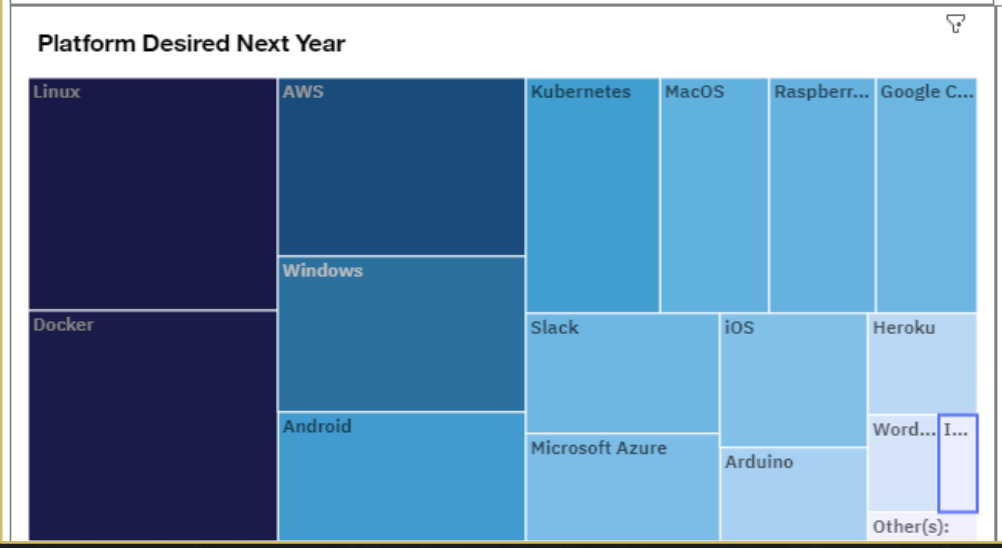
SKILLS NETWORK

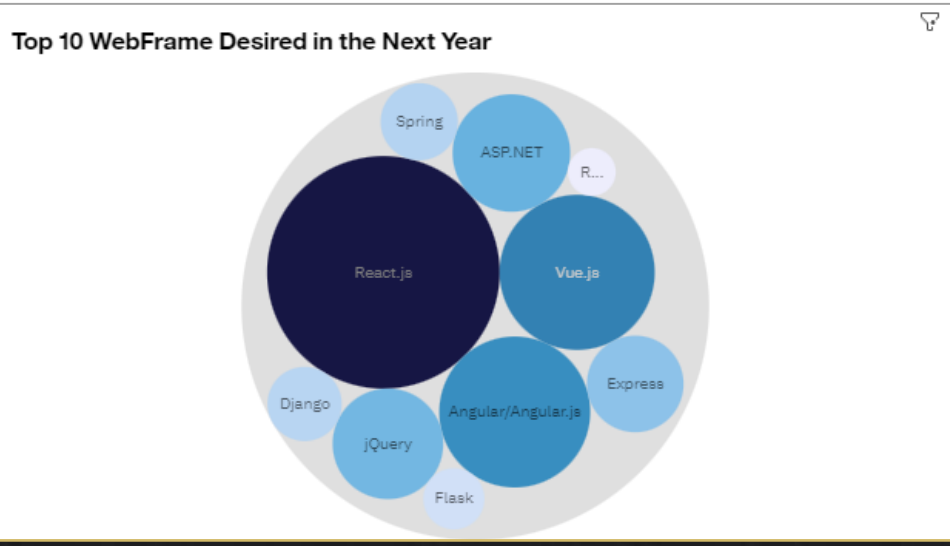
DASHBOARD TAB 2

Current Technology Usage | Future Technology Trend | Demographics





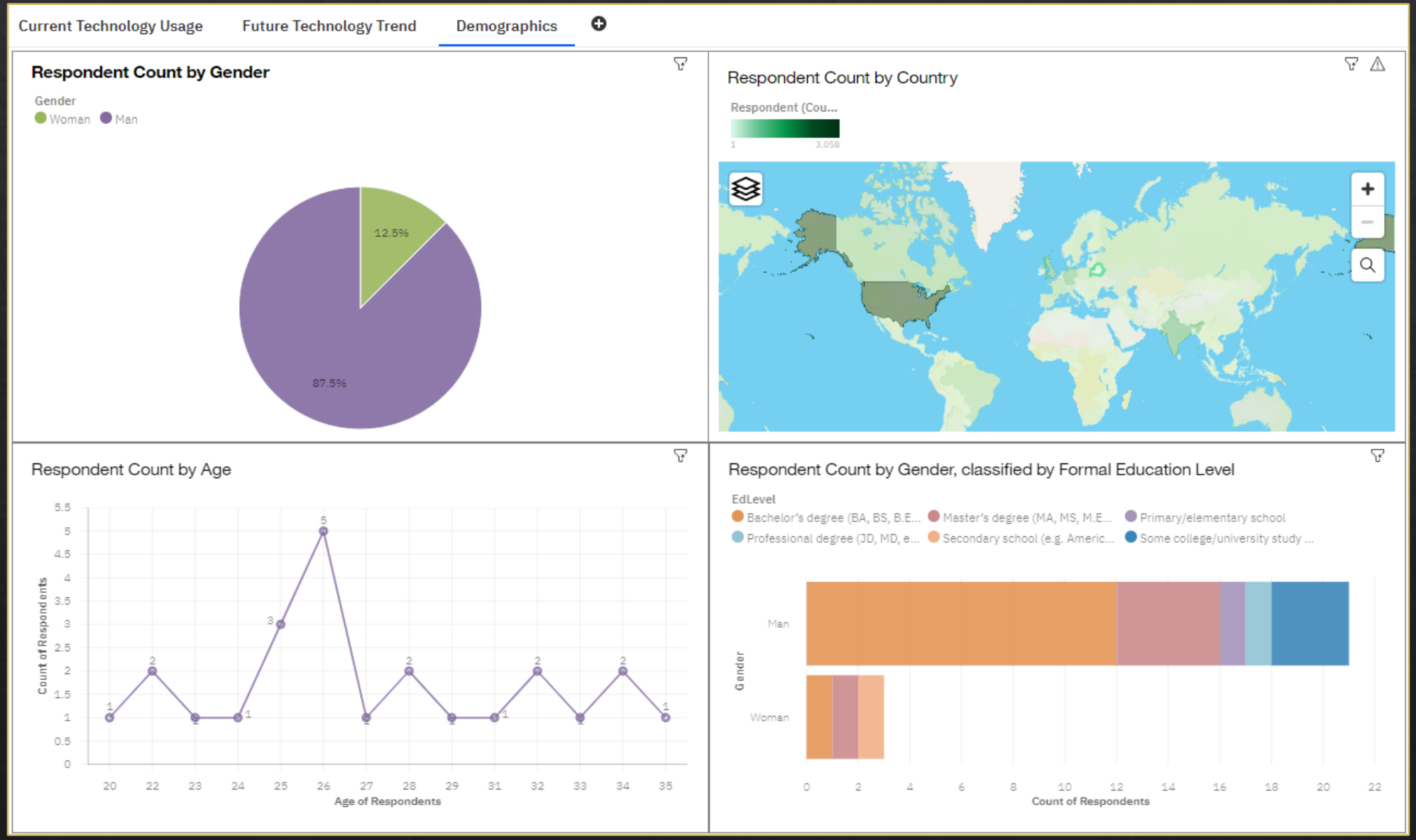




IBM Developer

SKILLS NETWORK

DASHBOARD TAB 3



IBM Developer

SKILLS NETWORK

DISCUSSION

- The findings of this analysis provide a number of insights. Some important questions to consider are:
  - Which developer technologies are in top demand?
  - Which technologies should prospective developers and data professionals be learning?
  - Which technologies should educators place more emphasis on in upcoming years?
  - Which technologies should businesses and organisations be emphasizing for their projects and operational requirements?
  - What is the developer demographic like?
  - Is there a gender representation gap?

OVERALL FINDINGS & IMPLICATIONS

Findings

- High usage and interest in Javascript, HTML/CSS remain in high usage. There's also increasing interest in Typescript
- Increasing interest in Python
- High usage and interest in SQL. MySQL had the highest use in 2019, but PostgreSQL is improving and was the most desired database for the next year
- NoSQL database is gaining interest, of which MongoDB was the most used in 2019 and desired for the next year
- There is a significant gender balance gap (in favour of men)

Implications

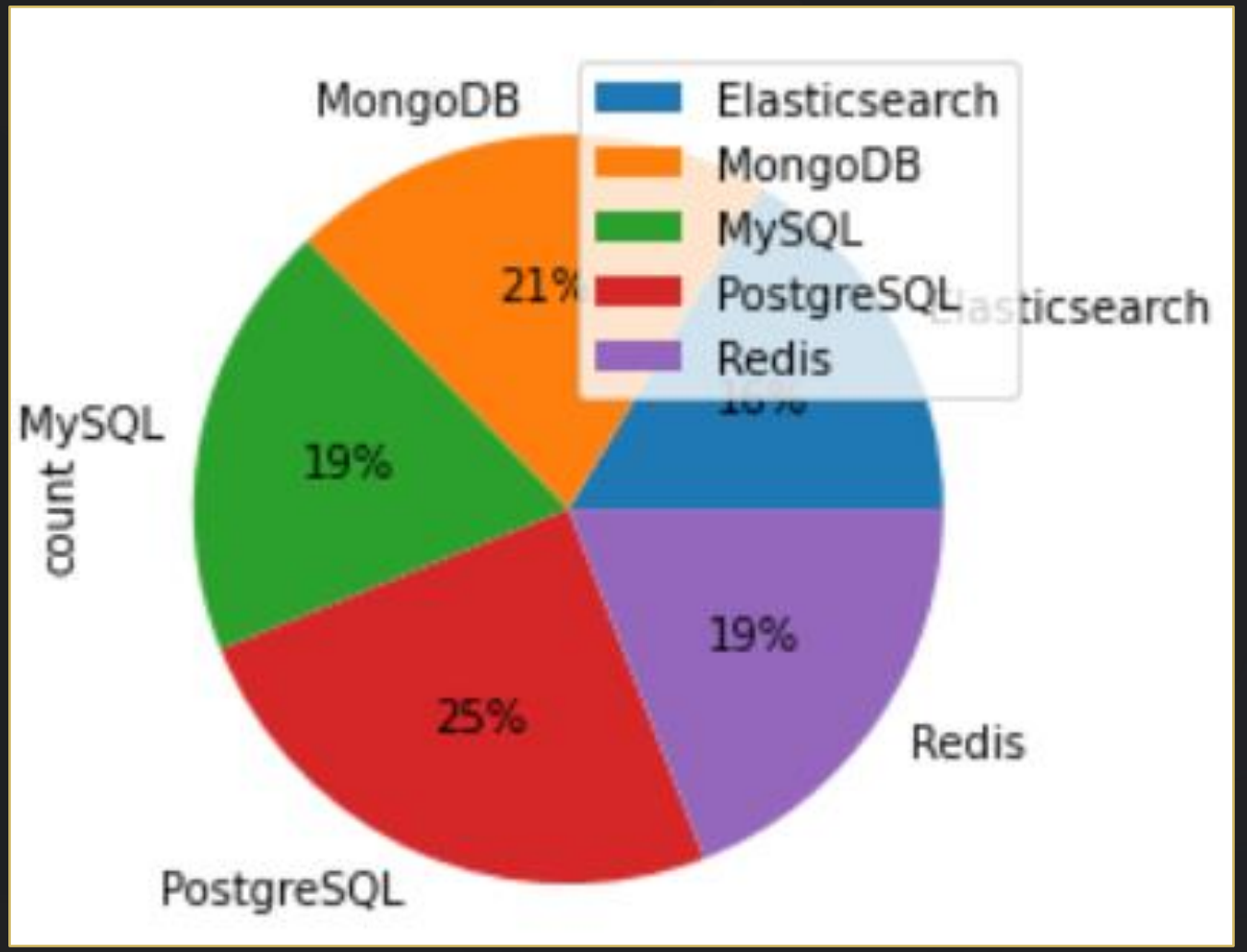
- Web development is in high demand. Current and prospective developers could consider picking up Typescript as well as Javascript and HTML/CSS
- With the growing need to handle big data and perform AI and ML work, data professionals should continue to improve SQL competence but also enhance skills in NoSQL and Python
- Businesses should adapt to changing technology preferences, particularly in terms of project types, talent acquisition and development.
- With the increasing popularity of these languages and applications it would be an opportune time to adjust policies which encourage female engagement in the industry

CONCLUSION

- The findings yielded various insights into technologies most used and desired by developers, including the developer demographic.
- These insights should be particularly relevant for current and prospective developers aiming to remain competitive, businesses aiming to upskill their talent, educators and policy makers aiming to address gender and economic issues.
- Businesses could look to these results as a way of steering their use of platforms and projects to align with the trends seen here. This could potentially avoid issues with technology projects affected by lack of developer resources.
- These same businesses would do well to initiate and/or nurture the development of employees using the common and desired languages and platforms.

APPENDIX

Top 5 Programming Languages used by Respondents

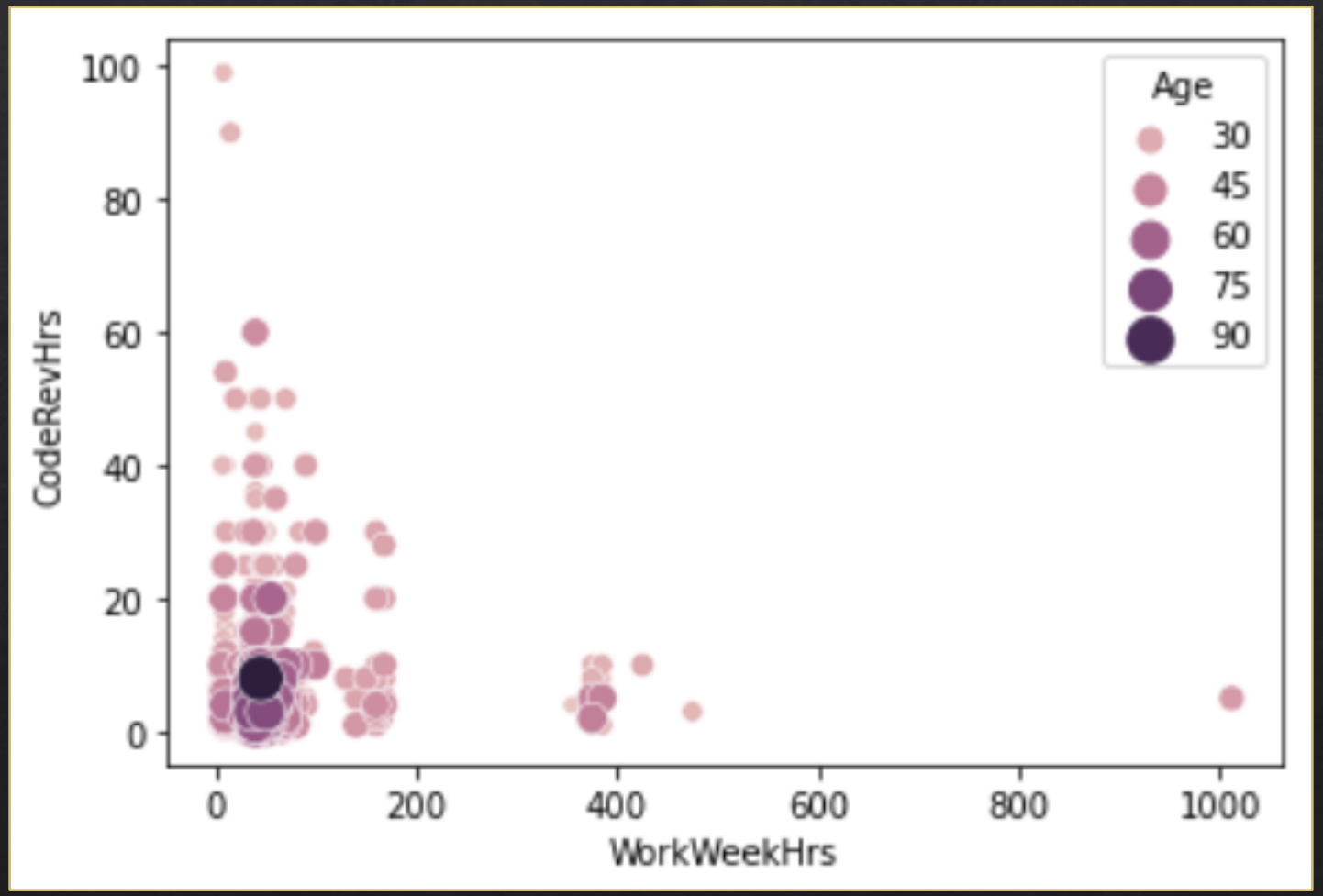


IBM Developer

SKILLS NETWORK

APPENDIX

Distribution of Hours Worked & Code Revision Hours, based on Age

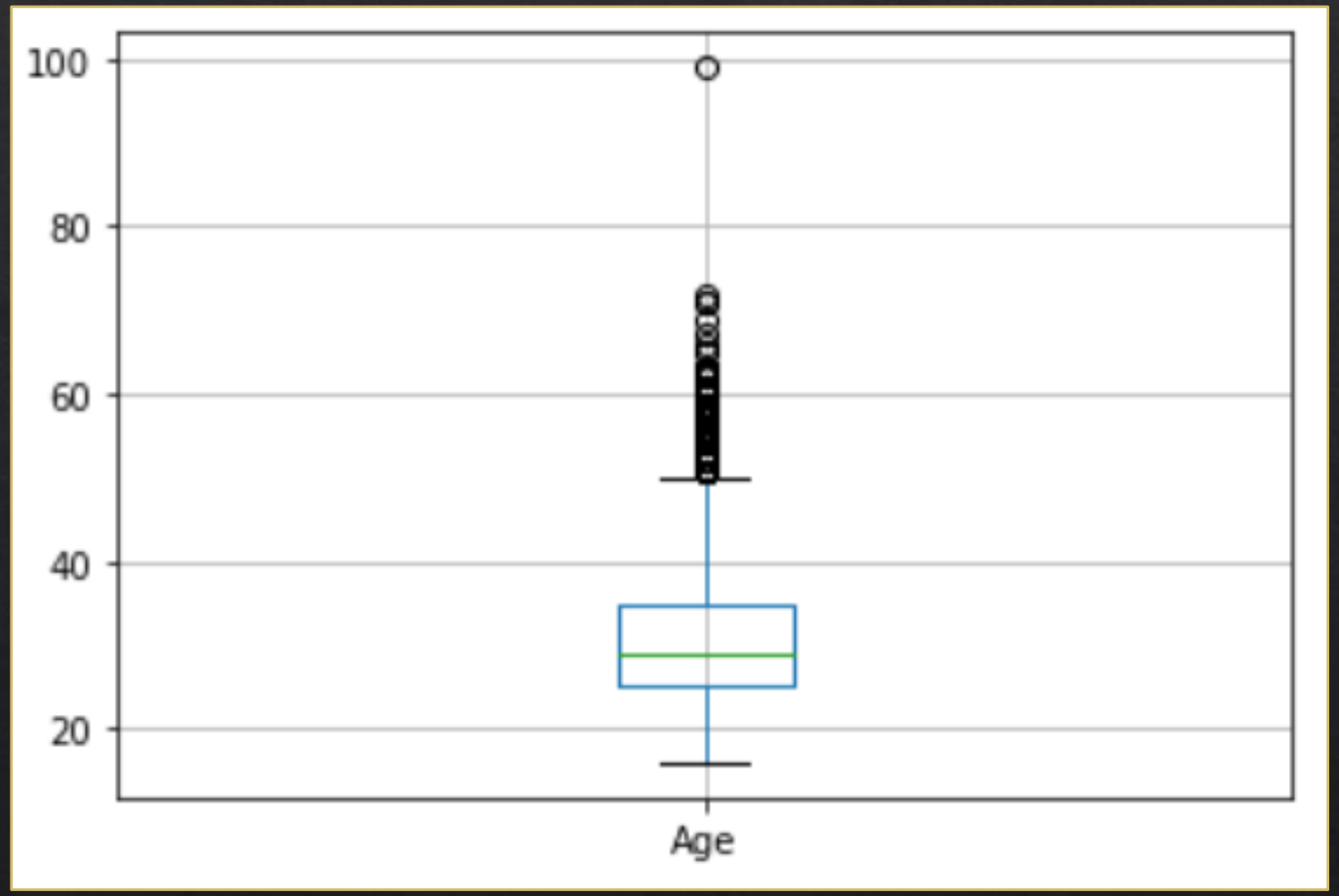


IBM Developer

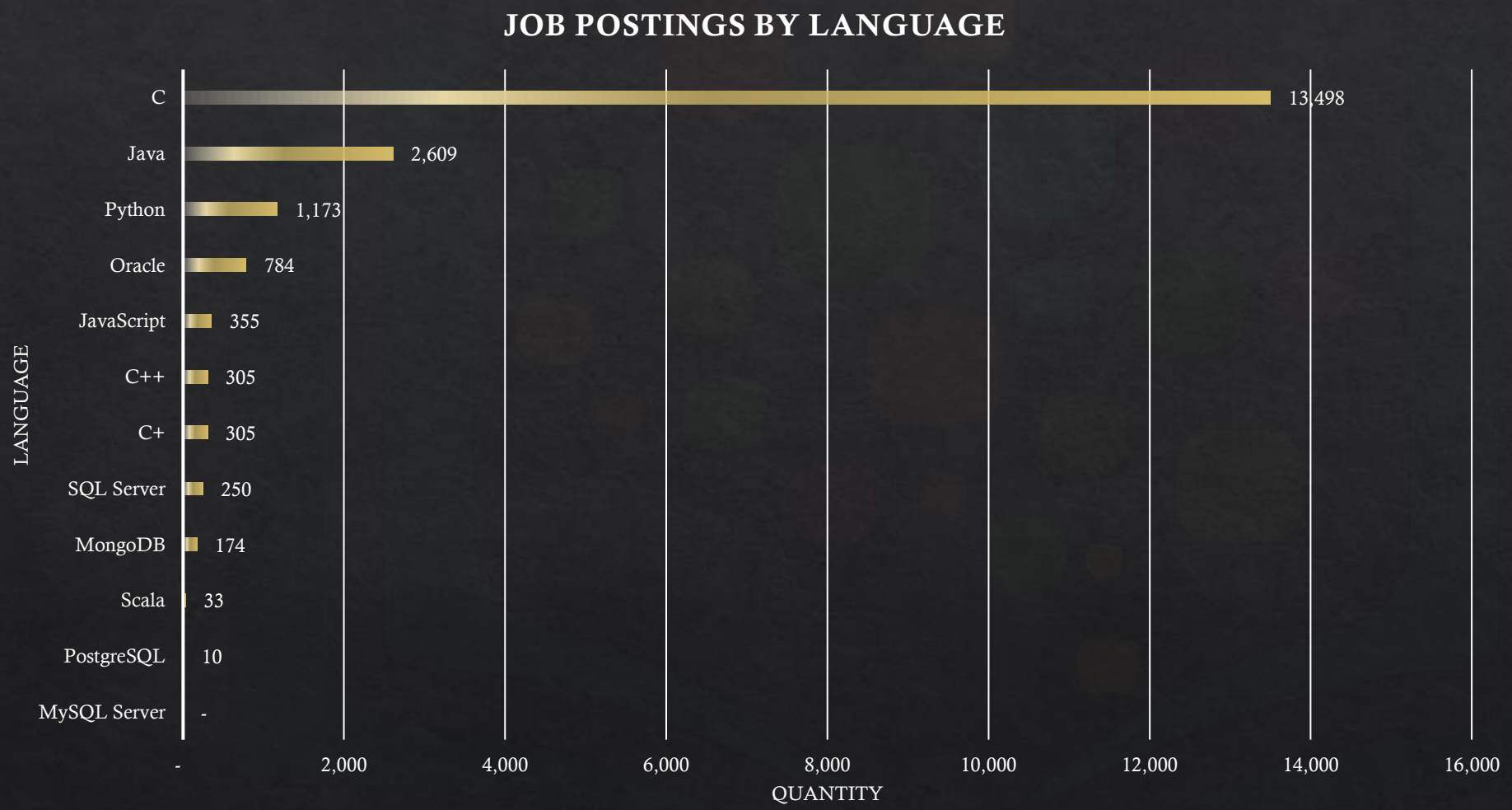
SKILLS NETWORK

APPENDIX

Age Distribution of Respondents



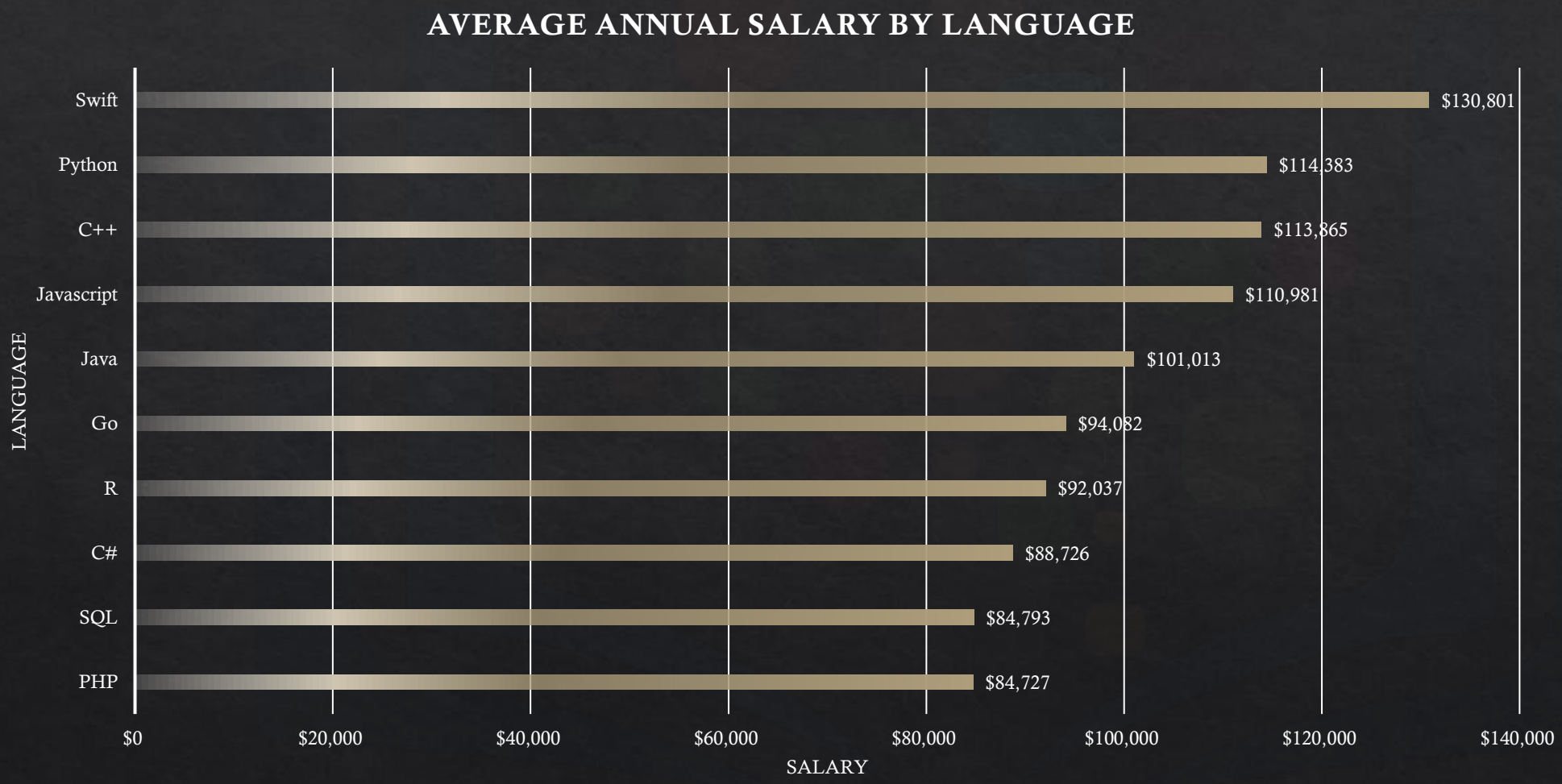
JOB POSTINGS



IBM Developer

SKILLS NETWORK

POPULAR LANGUAGES



IBM Developer

SKILLS NETWORK